Effect of the number of amplitude-compression channels and compression speed on speech recognition by listeners with mild to moderate sensorineural hearing loss

Marina Salorio-Corbetto[a)] and Thomas Baer

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

Michael A. Stone

Division of Human Communication, Development & Hearing, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

Brian C. J. Moore

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

[a)] Author to whom correspondence should be addressed. Electronic mail: ms878@cam.ac.uk

Running heading: Multichannel amplitude compression

Abstract

The use of a large number of amplitude-compression channels in hearing aids has potential advantages, such as the ability to compensate for variations in loudness recruitment across frequency and to provide appropriate frequency-response shaping. However, sound quality and speech intelligibility could be adversely affected due to reduction of spectro-temporal contrast and distortion, especially when fast-acting compression is used. This study assessed the effect of the number of channels and compression speed on speech recognition when the multi-channel processing was used solely to implement amplitude compression, and not for frequency-response shaping. Computer-simulated hearing aids were used. The frequency-dependent insertion gains for speech with a level of 65 dB SPL were applied using a single filter before the signal was filtered into compression channels. Fast-acting (attack 10 ms, release 100 ms) or slow-acting (attack 50 ms, release 3000 ms) compression using 3, 6, 12, and 22 channels was applied subsequently. Twenty adults with sensorineural hearing loss were tested using a sentence recognition task with speech in 2- and 8-talker babble at three different signal-to-babble ratios (SBRs). The number of channels and compression speed had no significant effect on speech recognition, regardless of babble type or SBR.

I. INTRODUCTION

Loudness recruitment, the greater-than-normal rate of growth of loudness with increasing sound level (Steinberg and Gardner, 1937), is present in nearly all cases of cochlear hearing loss. Compensation for hearing loss with linear amplification, for which the gain does not vary with input level, involves a compromise between audibility and comfort; the audibility of low-level sounds must be sacrificed to prevent loudness discomfort from high-level sounds. By using amplitude compression, it is possible to make low-level sounds audible while keeping high-level sounds comfortable. For this reason, most if not all modern hearing aids incorporate some form of amplitude compression (Moore and Popelka, 2016). Some early studies comparing linear amplification to amplitude compression suggested an advantage of amplitude compression for the recognition of speech in quiet (Laurence *et al.*, 1983; Villchur, 1973), especially for soft speech (Moore *et al.*, 1992; Bustamante and Braida, 1987; Lippmann *et al.*, 1981; Kam and Wong, 1999), or when the speech level varied rapidly over time (Lippmann *et al.*, 1981). Some studies also showed advantages for the recognition of speech in background noise (Laurence *et al.*, 1983; Moore *et al.*, 1992; Yund and Buckles, 1995b; Villchur, 1973; Yund *et al.*, 1987). These early studies were mostly performed using a fixed – and often small – number of amplitude-compression channels. However, the number of channels in hearing aids has increased over the years, with 20 channels or even more being employed in some hearing aids. The channels are typically used *both* for shaping the overall frequency-gain characteristic and for applying frequency-dependent compression. Hence, any advantage gained from increasing the number of channels might be a consequence of more accurate and flexible achievement of either or both of these uses. The main objective of this study was to assess the effect of the number of channels on speech intelligibility when the channels were *not* used for shaping the frequency-gain characteristic. The effect of the speed of compression and the interaction between the number of channels and the speed of compression were also investigated.

The use of many channels gives greater flexibility in shaping the overall frequency response to match a target prescription, such as the frequency-dependent insertion gain (IG) required for speech with an input level of 65 dB SPL, which is often used as the target when verifying a fitting using real-ear measurements (Mueller *et al.*, 2013). This use of multiple channels is similar to that of a graphic equalizer, and adjustments to match a given target are usually made by changing the gains in individual channels. The use of many channels compensates for the frequency dependence of loudness recruitment, allows the amount of compression to vary smoothly across frequency, and allows gain variations in a given frequency region to be more independent of those in adjacent frequency regions. However, amplitude compression using many channels may have undesired side effects, such as reduction of spectral contrast (Plomp, 1988; Stone and Moore, 2008), and, if the compression is fast acting, reduction of amplitude modulation depth (Plomp, 1988; Stone and Moore, 2008) and reduction of the correlation of envelope fluctuations in different frequency regions (Stone and Moore, 2008). Additionally, fast compression may distort the envelope shape (Stone and Moore, 1992) and lead to abrupt changes in the envelope magnitude at the onsets and offsets of sounds, known as overshoot and undershoot (Verschuure *et al.*, 1996; Stone and Moore, 2008; Moore, 2012). The envelope distortion arising from compression usually increases with increasing number of compression channels and increasing compression speed (Nábělek, 1983; Plomp, 1994; Kates, 2010; Holube *et al.*, 2016). Fast compression has little effect on the temporal fine structure (TFS) of sounds unless the compression time constants are comparable to one period of the signal at the output of a given channel (Levitt *et al.*, 1987).

An additional consequence of the use of fast-acting compression is an effect called “cross-modulation” or “across-source modulation correlation” (ASMC) (Stone and Moore, 2007), whereby the envelopes of independent sound sources become partially correlated after compression. This could lead to perceptual fusion of the sound sources, impairing the intelligibility of speech in competing-talker backgrounds. Fast-acting compression can also decrease the correlation of the modulation patterns across frequency channels for a given source (within-source modulation correlation, WSMC) (Stone and Moore, 2008), which might impair the ability to perceptually group the frequency components originating from a given source (Bregman *et al.*, 1985; Moore and Shailer, 1992).

Previous research on the effect of the number of amplitude-compression channels and compression speed has given mixed results. Some studies have shown that, with fast-acting compression, increasing the number of channels has a negative effect on speech recognition or subjective preference. For example, van Dijkhuizen (1993, unpublished data presented by Plomp, 1994) found that, for hearing-impaired listeners, performance on a sentence-recognition task worsened with increasing number of channels with the compression ratio set to 2 or 4 in all channels. Bor *et al.* (2008), who used a compression ratio of 3 in all channels, found that performance in a vowel-identification task was significantly better with two channels than with a single channel, but performance decreased when the number of channels was increased from 8 to 16. A problem with both of these studies is that the amount of compression was not selected based on the hearing loss of the listeners. More recently, Alexander and Masterson (2015) used 4-, 8-, and 16-channel simulated hearing aids with compression selected individually for each listener using the DSLv5 method (Scollie *et al.*, 2005) and found that the best outcomes were achieved when using four or eight channels with a short release time, or eight channels with a long release time. However, the effects were very small. Holube *et al.* (2016) used simulated hearing aids to obtain sound-quality ratings for different numbers of channels and different compression speeds. One condition used compression ratios selected using the DSL(i/o) method (Cornelisse *et al.*, 1995) and other conditions used either higher or lower compression ratios than prescribed using DSL(i/o). The results for the five hearing-impaired listeners tested showed that fewer compression channels, longer release times and lower compression ratios were generally preferred.

Other studies have shown that speech recognition improves with increasing number of channels up to a certain number and then remains constant. For example, Yund and Buckles (1995c) used very fast compression, and set the compression in each channel so that a stimulus presented at 20 dB HL in that channel was just audible, while no gain was applied for an input level of 100 dB SPL. They found that the recognition of consonants in noise improved with increasing number of channels from 4 to 8 and then remained roughly constant up to 16 channels (albeit with a slight dip in performance with 12 channels).

Finally, some studies have shown no change in performance with increasing number of channels when the compression was appropriately fitted for each listener. For example, Crain and Yund (1995) reported detrimental effects of fast-acting multi-channel compression when using high compression ratios and large numbers of channels (up to 31) when the compression ratio was fixed across channels and listeners. However, when the compression ratio was adjusted to be appropriate for each listener at each frequency, no detrimental effect of increasing the number of channels was found. Using fast compression fitted individually for each listener, Moore *et al.* (1999) found a small but significant advantage of having eight amplitude-compression channels over one, two, or four, when the background noise was deeply amplitude modulated and had relatively large spectral notches, but no effect of channel number was found for other backgrounds (single speaker, modulated noise, and spectrally notched noise). Keidser and Grant (2001) used one-, two-, and four-channel hearing aids with fast-acting compression fitted using the NAL-NL1 procedure (Byrne *et al.*, 2001) in a paired-comparison task and found that, overall, across the different stimuli used, most listeners did not have a preference for a particular hearing aid. These authors also evaluated performance in a speech-in-noise recognition task and found no significant effect of the number of channels.

The lack of agreement across these studies about the effect of number of channels probably arises partly from the diverse ways that the compression was implemented and set. For example, in many studies the values of the compression ratio were much higher than would be used in clinical practice (Nábělek, 1983; Plomp, 1994; Bor *et al.*, 2008). This could account for the deleterious effects of increasing number of channels observed in those studies (Crain and Yund, 1995). Also, several studies used compression that was very fast acting, and probably faster than typically used in commercial hearing aids. In the present study, compression ratios were calculated on an individual basis using a gain prescription formula (Moore *et al.,* 2010b) and time constants were chosen to be representative of those used in real hearing aids.

In some of the studies described above the multi-channel processing was used to achieve the desired frequency response (Keidser and Grant, 2001; Alexander and Masterson, 2015; Yund and Buckles, 1995c; Crain and Yund, 1995). The same is true for most commercial hearing aids. The match between the target frequency response and the output of the hearing aid is usually better when a greater number of channels is employed (Woods *et al.*, 2006), at least for numbers of channels up to about 5. Thus, any effects of number of channels could partly reflect the accuracy with which the frequency response matched the target rather than the effect of the number of compression channels *per se*. To our knowledge, only a few studies are not affected by this confound. Nábělek (1983) used a single filter after amplitude compression had been applied in order to equate the desired frequency response across conditions. However, this study was limited in scope, since the number of compression channels was either one or three. Bor *et al.* (2008) did not apply frequency-response shaping to their stimuli but instead presented their stimuli at 92-94 dB SPL to ensure audibility over a wide frequency range. A disadvantage of this approach is that the spectral balance of the vowel stimuli used was not suited to the hearing losses of the listeners. Excessive low-frequency content could have increased any negative effects of amplitude compression. In the studies described earlier, Moore *et al.* (1999) and Holube *et al.* (2016) achieved their target frequency responses by applying linear gain using an equal number of bands for all of their test conditions.

In the study reported here, the multi-channel compression was implemented after frequency-dependent amplification had been applied. The number of channels was systematically varied from 3 to 22 and both fast-acting and slow-acting compression were used.

Psychoacoustic measures of the ability to use TFS information and auditory-filter width were obtained to explore their relationship with the effects of compression speed and number of channels. Moore (2008) and Stone *et al.* (2008) proposed that listeners with poor sensitivity to TFS would rely heavily on temporal envelope cues and might be especially susceptible to the envelope distortion produced by fast-acting compression. Such listeners might therefore perform better with slow-acting compression. On the other hand, listeners with good TFS processing abilities might be more tolerant of the envelope distortion produced by fast compression and might be able to benefit from the improved audibility of low-level portions of a target signal in the dips of a competing background produced by fast compression. Such listeners might therefore perform better with fast compression. These proposals received some weak support from a study of Moore and Sek (2016b) but not from a simulation study of Hopkins *et al.* (2012). The present experiment was intended to provide a further test of these proposals. It was also hypothesized that the ability to understand speech in background babble would worsen with decreasing sensitivity to TFS.

We did not have a specific hypothesis about how auditory-filter width might be related to the optimum number of compression channels (if there is an optimum), although it has been proposed that the width of the compression channels should be greater than the width of the auditory filters to avoid reduced contrast in the internal representation of the spectrum with increasing number of channels (Laurence *et al.*, 1983). We did hypothesize that wider auditory filters would be associated with poorer overall performance in the recognition of speech in babble.

II. METHOD

Each listener took part in a preliminary hearing evaluation consisting of pure-tone audiometry and the "threshold equalizing noise", TEN(HL), test (Moore *et al.*, 2004) for the detection of dead regions (DR) in the cochlea. These are regions with no or very few functioning inner hair cells, synapses, or neurons (Moore, 2001; Moore, 2004). Listeners who met the criteria described below subsequently took part in the main experimental task, a sentence-identification task. Additional hearing tests were carried out in order to explore factors that could underlie individual differences in the outcome of the main experimental task. These tests were: (1) The notched-noise test (Glasberg and Moore, 1990; Patterson, 1976) to determine the width of the auditory filters; (2) Two measures of the ability to use TFS information: the difference limen for frequency (DLF) and a measure of sensitivity to interaural phase, the TFS Adaptive frequency (TFS-AF) test (Füllgrabe *et al.*, 2017).

A. Listeners

Twenty-seven listeners were recruited from the laboratory database or by advertising in various buildings of the University of Cambridge and in some local General Practitioners' surgeries. They were selected to have sensorineural hearing loss, as indicated by air-bone gaps of 10 dB or less from 0.5 from 2 kHz. Air-bone gaps were allowed to be up to 25 dB at 4 kHz, since such gaps do not necessarily indicate a conductive hearing loss (Margolis *et al.*, 2013). Listeners were required to have a hearing loss of at least 25 dB HL at most test frequencies, with at least three hearing thresholds ≥ 35 dB Hearing Level (HL) for three test frequencies below 6 kHz, and a hearing threshold at 4 kHz ≤ 70 dB HL. Two listeners with DRs at three or more consecutive test frequencies in both ears were excluded, since extensive DRs have perceptual consequences (Vickers *et al.*, 2001; Baer *et al.*, 2002; Vinay and Moore, 2007) that would have complicated the interpretation of the outcomes of this study. Only one ear of each listener was tested. If no DRs were present, the test ear was the better-hearing ear. If DRs were detected only in one ear, the contralateral ear was tested. If restricted DRs were detected bilaterally, the ear with fewer positive outcomes in the TEN(HL) test was chosen. In cases where the test ear was not the better-hearing ear, the test ear had an air-conduction hearing threshold that was never more than 40 dB higher than the bone-conduction threshold at the same frequency for the non-test ear; this made “cross-hearing” unlikely, especially since “closed” headphones were used (see section II.B). Three listeners who had flat hearing losses with thresholds around 45-60 dB HL had to be excluded because they could not perform the speech recognition task. One listener withdrew from the study at an early stage. A further listener had to be excluded due to an error in the processing of the stimuli. Complete results were obtained for twenty listeners (twelve women). Their median age was 73 yrs (range 45-86 yrs). All listeners were native speakers of British English.

B. Apparatus and stimuli

All measurements took place in a double-walled sound-attenuating chamber. Pure-tone audiometry was carried out using a Grason-Stadler GSI-61 audiometer (Eden Prairie, MN) equipped with Telephonics TDH-50 headphones (Telephonics, Huntington, NY), Sennheiser

HDA200 headphones (Wedemark, Germany), and a Radioear B71 bone-conduction transducer (Eden Prairie, MN). The TEN(HL) test and the notched-noise test were performed using a Philips compact disc player type 753 (Philips, Amsterdam, Netherlands) connected to the audiometer and using the Telephonics TDH-50 headphones. Measurement of the DLFs and the TFS-AF test were carried out using software developed by Sek and Moore (2012) and Moore and Sek (2016b), installed in a Samsung laptop (Seoul, South Korea) equipped with an M-Audio Audiophile USB external soundcard (Cumberland, RI). Stimuli were presented via Sennheiser HDA200 headphones.

For the speech-recognition task, the stimuli were generated by the Samsung laptop and converted to analog form by the M-Audio Audiophile USB external soundcard. The output of the soundcard was routed to an Aphex HeadPod 454 headphone amplifier (Long Beach, CA) and attenuated by 17 dB with a custom-built manual attenuator. Presentation was monaural via a Sennheiser HDA200 headset. The stimuli were sentences from the STARR corpus (Boyle *et al.*, 2013), uttered by a male speaker. Sentences were mixed with either 2-talker babble, composed of a female and a male speaker, or an 8-talker babble composed of four female and four male speakers. The babble sounds were prepared from recordings made by Moore *et al.* (2008). The target speech and babble backgrounds were separately filtered to match their long-term average spectra to the long-term average spectrum of speech described by Moore *et al.* (2008). Next, the target speech was mixed with each type of babble at signal-to-babble ratios (SBRs) of –3, 0, or +3 dB. The speech level at the input to the simulated hearing aid was always 65 dB SPL, and the noise level was varied to achieve the desired SBR. The babble always started 1 second before the target speech and ended 0.5 seconds after it.

C. The simulated hearing aid

The speech in babble was processed using a computer-simulated hearing aid similar to that described by Moore *et al.* (2011). However, here the simulated hearing aid processed the waveform files on a sample-by-sample basis rather than using the overlap-add method that was employed earlier. The sample rate was 22050 Hz. All files were initially high-pass filtered at 50 Hz with a forward and backward pass of a third-order elliptic filter with 0.1-dB passband ripple and 30-dB stopband ripple. This removed any undesirable infrasonic components in the input file, which might otherwise bias level estimates. The aid applied the frequency-dependent insertion gains (IGs) recommended by the CAM2B procedure (Moore *et al.*, 2010b; Moore and Sek, 2016a) for speech with a level of 65 dB SPL at frequencies 0.125, 025, 0.5, 1, 2, 3, 4, 6, 8, and 10 kHz, using a single linear-phase finite impulse response (FIR) filter with 133 taps. The impulse response duration was 6.03 ms, corresponding to a delay of 3.02 ms, since the inherent delay produced by an FIR filter is one-half of the impulse-response duration. The gains at intermediate frequencies were specified by interpolation on a dB versus linear-frequency scale (the outcome would have been very similar if a log-frequency scale had been used, since the filter gain was specified at frequencies that were spaced by one-half octave or less). This filter is called hereafter the "insertion-gain filter". The frequency response of the headphone, as measured using KEMAR (Burkhard and Sachs, 1975), was allowed for when calculating the filter response, so as to achieve the correct IGs.

Multi-channel compression using 3, 6, 12, or 22 channels was applied subsequently. Thus, the channels were not used for frequency-response shaping. The channels had equal widths on the ERB_N-number scale (Glasberg and Moore, 1990). The channel filters were designed so that, when the channel outputs were combined, spectral ripple was minimal (less than ±1 dB). Filtering to create the channels was performed using linear-phase FIR filters with a maximum impulse-response duration of 14.74 ms (corresponding to a delay of 7.37 ms). The length of the FIR filters varied across channels, which introduced an across-frequency delay. This was removed by use of delay lines so as to time-align the channel signals. An additional delay of 2.5 ms was needed to implement the attack of the compression system, so the overall delay produced by the processing was 3.02 + 7.37 + 2.5 = 12.89 ms. This is close to but slightly longer than would be desired for a practical application (Stone and Moore, 1999; Stone and Moore, 2002). The skirts of the filter frequency responses were nearly identical when plotted on a logarithmic frequency scale. The lowest channel was created using a low-pass filter, while the remaining channels were created using band-pass filters. The maximum relative response in the stopband of the filters was –60 dB for the 22-channel implementation, decreasing to –68 dB for the 3-channel implementation. The effective frequency range of the simulated hearing aid was 50 to 9900 Hz.

As noted above, the channel filters had similar slopes regardless of channel number. This had two consequences. Firstly, the level in a given channel prior to the application of compression deviated more from the level produced by an ideal rectangular filter as the number of channels increased. Secondly, the overall output level after combining the channels increased with increasing number of channels. These two effects were compensated using a two-stage calibration process which was conducted without any compression processing.

The first stage made use of a noise with a spectrum corresponding to a stylized version of the long-term average speech spectrum (LTASS) described by Moore *et al.* (2008). The spectrum level of the noise was flat from 100 to 500 Hz, decreased by 7.5 dB/octave up to 8000 Hz, and then decreased by 13 dB/octave up to the Nyquist frequency of 11025 Hz. The insertion-gain filter was applied to a 5-s sample of the noise and the power spectral density of the output was calculated using the Welch method. This in turn was used to calculate the “ideal” level at the output of each channel filter, assuming rectangular filters. The same noise sample was used as input to the insertion-gain filter and the channel filters, and the output level of each channel filter was measured. The differences in level between the ideal and actual filter outputs were used to “correct” (reduce) the output level of each channel so that it corresponded to the ideal output level. The application of this correction meant that channel output levels could be used to accurately calculate the required gains for each channel in response to other input signals, when compression processing was applied.

In the second stage of calibration, the difference was calculated between the idealized channel output levels and the corresponding levels over the same frequency range when the outputs of all channels were summed. The difference represents the increase in level due to the overlap of the bandpass filters, which increased with increasing number of channels. The difference was used to reduce the level of the signal at the output of each channel, after compression processing but prior to summation of the channel signals.

The net effect of these two stages of correction was that, for a fixed selection of frequency-dependent insertion gains, the spectral shape of the output in response to noise with the LTASS was almost independent of the number of channels for frequencies up to 8000 Hz. There was a very small effect of the number of channels (less than ± 1 dB) at 9900 Hz, which corresponded to the upper edge of the highest channel.

Because of our use of a simulated hearing aid and the corrections described above, the target IGs could be achieved accurately over a wide range of input levels and center frequencies. The main source of deviations from the target gains was the variability of the response of the headphones across individual ears. Measurements conducted in our laboratory obtained using an Etymotic Research ER10C probe microphone indicated that the standard deviation of the response across ears was typically less than ±3 dB for frequencies up to 4 kHz, and less than ±5 dB for frequencies from 4 to 10 kHz.

The compression had a main stage, which controlled the gains of the simulated aid most of the time, and a second limiting stage that was intended to prevent loudness discomfort that might otherwise be produced by brief high-level peaks in the input signal. The compression ratio for the main stage was participant-dependent. The compression ratio for the second stage was 100. The main stage was either fast acting (attack 10 ms, release 100 ms) or slow acting (attack 50 ms, release 3000 ms); attack and release times were defined according to (ANSI, 2014). The compression threshold for each channel for the main stage was 15 dB below the root-mean-square (RMS) level within that channel when the input was a signal with the LTASS and a level of 65 dB SPL. Since a speech input level of 65 dB SPL was used in the experiment, this means that the speech signal was compressed over its entire effective dynamic range.

For each channel, the compression threshold for the second stage was usually set 15 dB above the RMS output level of that channel calculated from the first stage. The second stage was activated only rarely (usually 0.1% of the time or less), since short-term peaks in the level of speech only rarely have a level more than 10 dB above the RMS level in a given frequency region (Moore *et al.*, 2008). Rarely, the second stage was activated more than 0.1% of the time in a few channels, in which case the compression threshold for the second stage was increased by up to 5 dB. The attack time for the second stage was 6 ms for the lowest channel and decreased linearly to 2 ms for the highest channel. The release time was 80 ms for the lowest channel and decreased linearly with increasing channel number to 40 ms for the highest channel.

The hearing aid gains and compression ratios for the main stage were set separately for each listener using the CAM2B procedure (Moore *et al.*, 2010b; Moore and Sek, 2016a), based on the audiogram of the test ear. When fast compression was used, the compression ratio was limited to 3, since there is evidence that compression ratios above 3 have deleterious effects when fast compression is used (Neuman *et al.*, 1998; Keidser *et al.*, 2011). We believe that this is representative of what happens in clinical practice. The compression ratio was limited to 10 when slow compression was used. Figures 1 and 2 show the compression ratios used for each compression speed and number of channels.

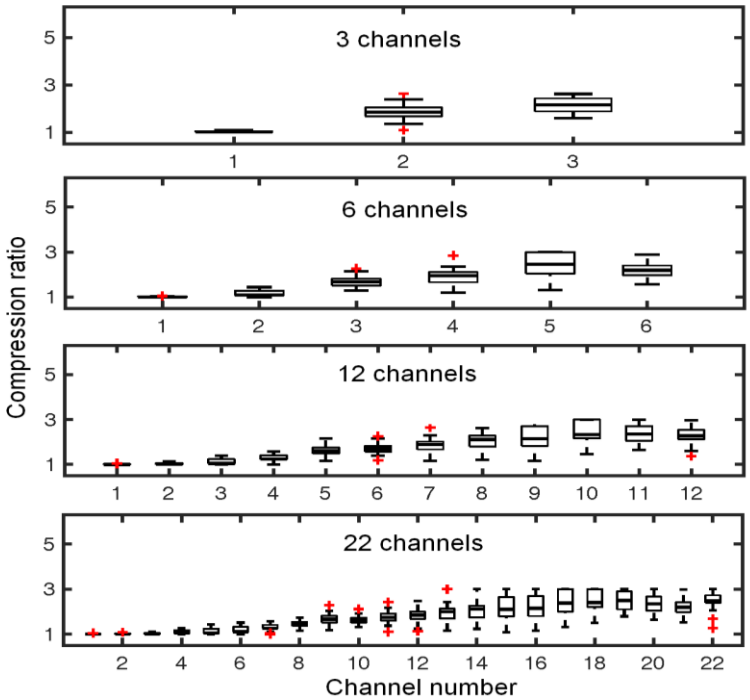

FIG. 1. Compression ratios for each channel when fast compression was used. The lower edge of each box indicates the 25th percentile and the upper edge of each box indicates the 75th percentile. Whiskers indicate the extreme data points that were not considered as outliers. Outliers are indicated by crosses (color online). The median is indicated by a thick line.

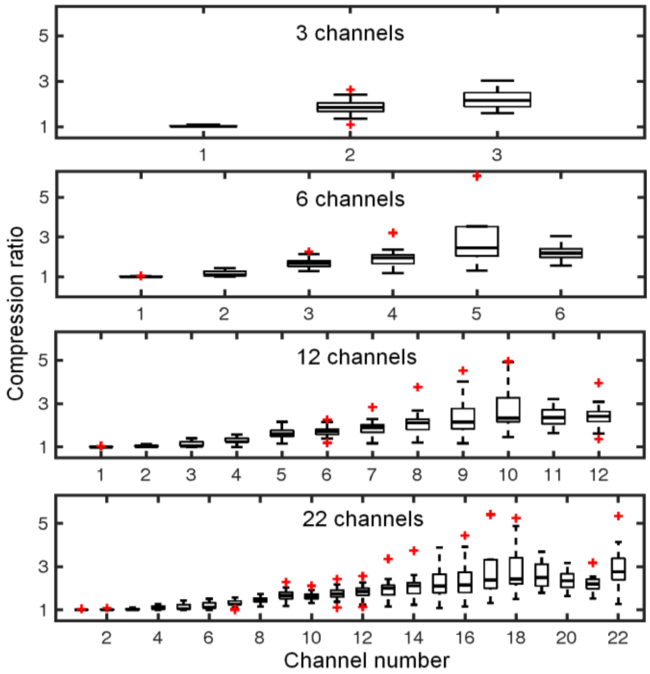

FIG. 2. As Fig. 1 but for slow compression.

To quantify the effect of the compressors on the envelopes of typical speech stimuli, we used the measure "fidelity of envelope shape" (FES) described by Stone and Moore (2007). The FES varies from 1 if the envelopes in different frequency channels are perfectly preserved to 0 if envelope information is completely destroyed. FES values were calculated for the 22-channel compression systems, for which the effects were expected to be greatest. The compression ratios were set to the median values shown in Figs 1 and 2. The input signal was a one-minute sample of a single female talker mixed at 0 dB SBR with either two or eight talkers (with an even number of males and females in each). For the two-talker background the FES was 0.7244 for the fast compressor and 0.7447 for the slow compressor. For the eight-talker background, the FES was 0.6861 for the fast compressor and 0.7078 for the slow compressor. Thus, both compressors altered the envelope shapes of the signals, and the alteration was somewhat greater for the fast than for the slow compressor.

In summary, there were 48 conditions for the speech intelligibility test, given by four numbers of channels (3, 6, 12, or 22), two compression speeds (fast or slow), two backgrounds (2- or 8- talker babble), and three SBRs (–3, 0, or +3).

D. Procedure

1. Pure-tone audiometry

Pure-tone air- and bone-conduction audiometry was performed using the procedure recommended by the British Society of Audiology (2011). For air conduction, thresholds for octave and semi-octave frequencies between 0.125 and 10 kHz were obtained. For bone conduction, octave frequencies from 0.25 to 4 kHz were used.

2. TEN(HL) Test for the detection of dead regions in the cochlea

To detect any DRs, the TEN(HL) test (Moore *et al.*, 2004) was conducted for frequencies between 0.5 and 4 kHz. A DR was deemed to be present at a specific test frequency when the masked threshold of the test tone in the TEN was 10 dB or more above the hearing threshold in quiet and 10 dB or more above the nominal level of the TEN (Moore, 2004). A difference of 8 dB between the masked threshold and the level of the TEN was considered as 'inconclusive' (Moore, 2004), and in such cases the test was repeated using a higher level of the TEN, if possible. In most cases, the nominal level of the TEN was at least 10 dB above the hearing threshold for the test frequency in quiet. In two cases (L8 at 4 kHz and L26 at 3 kHz), the nominal level of the TEN was set to be the same as the hearing threshold for the tone or only slightly above this level, to avoid loudness discomfort. For the same reason, in one case (L26 at 3 kHz), the nominal level of the TEN was slightly below the threshold in quiet.

3. Recognition of sentences

Half the listeners completed the test first with the 2-talker background, and half with the 8-talker background. The two types of backgrounds were used on different days. Following the presentation of each sentence, the listener was asked to write down each word they heard. They were encouraged to guess when they were uncertain. Each sentence contained five key words, and only key words were scored. Each session started with eight sentences in quiet for familiarization and 24 practice sentences (one for each condition for the type of babble being used on that day). Conditions were tested in blocks, first in a random order and then in the reverse order for each listener, to compensate for the effects of learning and fatigue. Each block used six randomly chosen sentences. The scores for the first two sentences in the block were discarded, as pilot work showed that listeners needed to get used to the listening condition to provide stable answers. The remaining four sentences in each block contained 20 keywords. Words identified correctly for each condition were added across two blocks, and the percent correct was calculated based on 40 keywords. For two of the listeners, one of the blocks for one condition in each case had to be discarded due to an error in the processing of the stimulus. The scores for these conditions were computed based on 20 keywords only. The resulting score was transformed into rationalised arcsine units (RAUs) using the formulae recommended by Sherbecoe and Studebaker (2004), which include a correction to account for the number of items, before performing statistical analysis. The total test time for each listener was 8 to 12 hours.

4. Determination of the auditory filter width

An abbreviated version of the notched-noise test described by Glasberg and Moore (1990) and based on work conducted by Patterson (1976) was used to derive the shapes of the auditory filters at center frequencies of 0.5, 1, and 2 kHz. The level of the pure-tone signal was fixed and the level of the notched noise was varied to determine the level at which the signal was just detectable. The notch in the noise was symmetric around the signal frequency. The notch widths, expressed as the deviation of each edge of the notch from the signal frequency divided by the signal frequency, were 0, 0.1, 0.2, and 0.3. The test tones were pulsed on and off. Each pulse had a duration of 200 ms including 20-ms raised-cosine ramps. The interval between pulses was 200 ms.

For each signal frequency, the absolute threshold in quiet was determined using a 2-dB step size with a method similar to that used for pure-tone audiometry (British Society of Audiology, 2011). Then, the signal level was fixed at 10 dB Sensation Level (SL). The level of the noise required for the signal to be just detectable was then determined for each notch width, using a final step size of 2 dB. Outcomes were analysed using the method described by Glasberg and Moore (1990) and using the software developed by them.

5. Difference limen for frequency (DLF)

Although the DLF could be based on the use of TFS information or place information (based on shifts in the excitation pattern), it has been argued that, at least for listeners with normal hearing, the TFS mechanism is dominant at 2 kHz (Moore, 1973; Goldstein and Srulovicz, 1977). Here, following Moore and Sek (2016b) we take the DLF at 2 kHz as an indirect measure of the ability to use TFS information at that frequency.

The DLF at 2 kHz for the test ear was measured with a procedure similar to that used by Moore and Ernst (2012) and Moore and Sek (2016b). First, the absolute threshold for a 2-kHz pure tone was determined using a three-interval forced-choice procedure using a 2-down 1-up tracking procedure to estimate the 71%-correct point on the psychometric function. Then, the DLF was measured at 20 dB SL. A two-interval forced-choice task was used where one of the intervals contained four tones whose frequency was 2 kHz and the other interval contained four tones whose frequencies alternated between 2 kHz and 2 kHz +Δf. Each tone lasted for 400 ms including 20-ms raised-cosine ramps, and the gap between tones within an interval was 100 ms. The silent gap between interval 1 and 2 was 400 ms. Listeners were asked to choose the interval in which there was a fluctuation in pitch. To reduce the usefulness of loudness cues induced by the frequency changes, the level of every tone was roved independently over the range ± 4 dB. The value of Δf was varied across trials using a 2-down 1-up procedure. Eight reversals were obtained. Δf was changed by a factor of 1.25^3 until the first reversal occurred, by a factor of 1.25^2 until the second reversal occurred, and by a factor of 1.25 thereafter. The DLF was estimated as the geometric mean of the values of Δf at the last six reversals. Five estimates of the DLF were obtained, and the final estimate was taken as the geometric mean of the last four.

6. The TFS-AF test

The TFS-AF test (Füllgrabe *et al.*, 2017) was used to estimate the highest frequency at which an interaural phase difference (IPD) of 180° could be distinguished from a reference value of 0°. This test used a two-interval forced choice paradigm with stimulus durations and structure the same as described above for the DLF task. One interval contained four successive tones with an IPD of zero and the other interval contained four tones whose IPD alternated between 0° and 180°. Listeners were asked to indicate which of the two intervals contained a sequence of tones that appeared to move within the head. The level of the tones was 30 dB SL. The required levels were calculated from the audiogram for each ear of each listener.

Listeners performed a training task using a single interval where they could control whether the four tones all had an IPD of 0° (‘Not moving’) or whether two of them had an IPD of 180° (‘Moving’). The frequencies of the stimuli used for this task were 200, 280, and 336 Hz. After the listeners confirmed that they could hear the difference across conditions, testing proper began. For testing proper, the frequency of the tones was initially 200 Hz. The frequency was adaptively varied using a 2-up, 1-down rule. The frequency was changed by a factor of 1.4 until the first reversal, by a factor of 1.2 until the next reversal, and by a factor of 1.1 thereafter. Each run was terminated after eight reversals. The threshold estimate was defined as the geometric mean of the frequencies at the last six reversals. Four estimates were obtained, and the final estimate was taken as the geometric mean of the last three.

Although the TFS-AF test assesses binaural sensitivity to TFS, this sensitivity is correlated with a measure of monaural sensitivity to TFS (Moore *et al.*, 2012), so it seemed reasonable to use the TFS-AF test here, even though our measures of speech intelligibility were all obtained under monaural listening conditions.

III. RESULTS

A. Pure-tone audiometry

Most listeners had gently sloping hearing losses. Table 1 shows the hearing thresholds of the ears tested and also gives information about previous experience with hearing aids.

TABLE I. Age (yr), hearing-aid (HA) use, and audiometric thresholds for the test ear of each listener. ‘Y’ indicates that the listener wore hearing aids, ‘N’ indicates that they did not. Columns four to 15 show audiometric thresholds (in dB HL) for each test frequency (in kHz). The mean, standard deviation (SD) and median (Md.) are shown.

ID	Age	HA	0.125	0.25	0.5	0.75	1	1.5	2	3	4	6	8	10
5	85	Y	15	20	40	35	35	35	40	35	40	70	60	75
8	72	N	15	20	35	40	45	45	60	65	70	65	70	75
10	80	Y	15	20	45	45	40	35	35	40	50	50	55	75
12	76	Y	15	25	30	25	35	35	45	50	45	50	50	80
13	79	Y	20	25	45	30	40	50	45	50	40	45	50	75
14	45	N	5	5	20	25	30	40	45	45	50	30	30	45
15	73	Y	10	15	20	25	25	30	35	30	45	45	60	60
17	65	Y	10	10	25	35	30	30	30	30	45	60	65	70
19	86	Y	15	10	20	25	35	45	45	45	55	55	70	85
20	78	Y	25	35	40	45	40	40	45	55	70	60	70	65
21	59	Y	10	5	15	25	25	20	30	40	45	35	25	20
23	70	Y	10	5	10	15	10	35	45	50	40	45	50	65
25	66	Y	0	0	-5	10	20	45	50	55	60	60	55	60
26	79	Y	15	15	10	5	30	5	15	60	70	65	70	80
27	70	Y	20	20	25	30	30	40	50	45	30	40	55	60
28	83	N	15	5	10	20	15	30	45	45	60	65	85	85
30	73	Y	10	25	35	35	30	30	25	5	20	40	65	70
31	71	N	25	25	20	25	25	20	10	30	50	55	60	50
32	67	Y	10	15	20	25	30	35	35	45	45	50	65	60
34	75	Y	15	20	30	25	20	35	45	60	60	55	60	60
Mean	72.6		13.8	16.0	24.5	27.3	29.5	34.0	38.8	44.0	49.5	52.0	58.5	65.8
SD	9.5		6.0	9.1	13	10	8.9	10	12	14	13	11	14	15
Md.	73.0		15.0	17.5	22.5	25.0	30.0	35.0	45.0	45.0	47.5	52.5	60.0	67.5

B. TEN(HL) test

DRs at 4 kHz were identified for L8 and L20, both of whom had audiometric thresholds of 70 dB HL at 4 kHz; DRs are common when the audiometric threshold is 70 dB HL or more (Aazh and Moore, 2007). The results were inconclusive at 3 and 4 kHz for L26 and at 4 kHz for L25.

C. Recognition of sentences

Figure 3 shows the outcome of the sentence-recognition task. Scores did not vary clearly with number of channels or compression speed. A four-way within-subjects analysis of variance (ANOVA) with factors number of channels, compression speed, type of babble, and SBR showed no significant effect of compression speed ($F(1, 19) = 0.05$, $p = 0.82$) or type of babble ($F(1, 19) = 0.37$, $p = 0.55$). The effect of number of channels just failed to reach the 0.05 level of significance ($F(3, 57) = 2.47$, $p = 0.071$). There was a significant effect of SBR, as expected, intelligibility increasing with increasing SBR ($F(2, 38) = 1387.0$, $p<0.001$, partial $\eta^2 = 0.99$). There was a significant interaction between SBR and type of babble ($F(2, 38) = 78.2$, $p<0.001$, partial $\eta^2 = 0.81$). For the lowest SBR, performance was better for the 2-talker babble, while for the highest SBR, performance was slightly better for the 8-talker babble. No other interactions were significant.

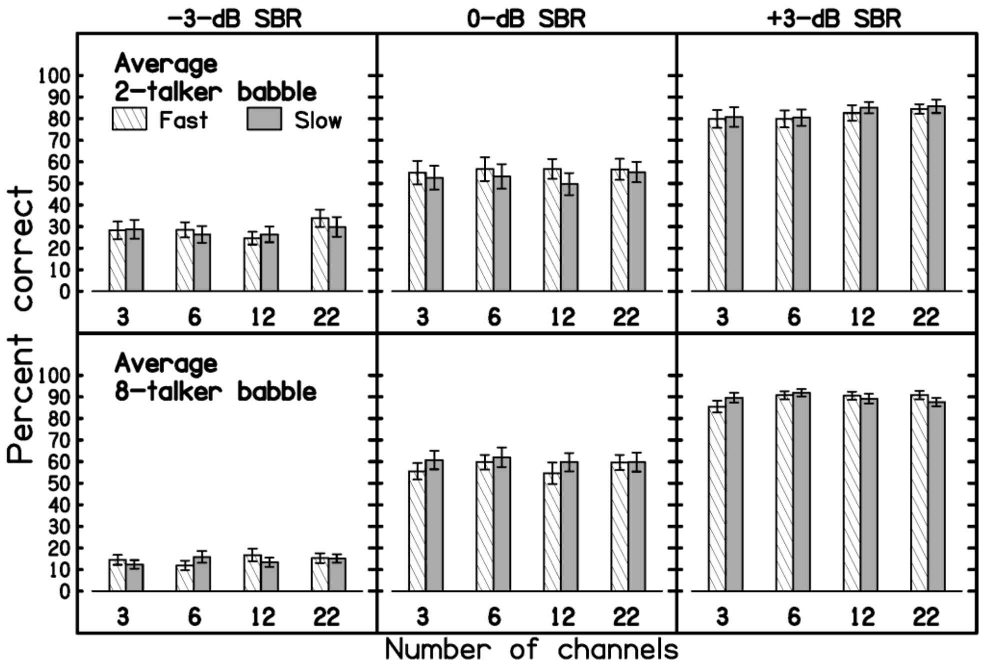

FIG. 3. Average scores for speech identification. Each column shows results for one SBR. The top and bottom panels show scores with the 2-talker and 8-talker backgrounds, respectively. The shading of the bars indicates compression speed, as indicated in the key. Error bars show ± 1 standard error.

D. Auditory filter width

The interpretation of the estimates of auditory filter bandwidth is complicated by the fact that the signal level in dB SPL varied across listeners and across frequencies, since the auditory filter tends to broaden with increasing level (Moore and Glasberg, 1987; Rosen *et al.*, 1992). To allow for this, we followed the procedure of Hopkins and Moore (2011), and expressed the bandwidth estimates relative to what would be expected for normal-hearing listeners tested at the same signal level. The auditory filters were normal or only slightly wider than normal (widening factor, i.e. bandwidth relative to the normal value, < 1.2) for eleven listeners at 0.5 kHz, thirteen listeners at 1 kHz, and three listeners at 2 kHz. There were two listeners, L30 and L31, for whom all three measured bandwidths were normal or close to normal. For the center frequency of 0.5 kHz, the widening factor ranged from 0.83 to 4.26, with a median of 1.14. For the center frequency of 1 kHz, the widening factor ranged from 0.82 to 2.16, with a median of 1.10. For the center frequency of 2 kHz, the widening factor ranged from 0.93 to 3.25, with a median of 1.73. The greater widening at 2 kHz is consistent with the sloping hearing losses of the listeners (Moore, 2007; Glasberg and Moore, 1986). To obtain an overall measure of frequency selectivity, the geometric mean widening factor across the three frequencies was calculated for each listener and used for subsequent statistical analysis. In what follows, this is termed ‘overall filter widening’. The overall filter widening ranged from 0.96 to 2.16, with a median of 1.30.

E. Difference limen for frequency

The DLF for L5 was not measured due to limited availability of this listener. The DLFs for the remaining 19 listeners ranged from 6.2 to 50.7 Hz, with a median of 19.3 Hz. Expressed as a percentage of the center frequency, the DLFs ranged from 0.3 to 2.5%, with a median of 1%. These values are higher than those for normal-hearing listeners obtained using the same task (Moore and Ernst, 2012), but are similar to those reported by Ernst and Moore (2013) and Moore and Sek (2016a) for hearing-impaired listeners.

F. TFS-AF test

One listener (L20) had a very large hearing loss in the non-test ear and therefore the TFS-AF test was not performed. L19 was not available for testing. The thresholds for the remaining 18 listeners ranged from 604 to 1283 Hz with a median of 953 Hz. The thresholds are comparable to those found previously for older listeners with hearing loss (Moore and Sek, 2016b; Füllgrabe and Moore, 2017).

G. Inter-relationship between measures

To assess the hypotheses described in the introduction, relationships between variables were explored by computing Pearson correlations. The DLFs were log-transformed to make their distribution more normal.

1. Psychoacoustic measures and overall speech intelligibility

To reduce the effect of random errors of measurement, an overall measure of intelligibility was calculated for each listener by averaging the scores in RAUs across all backgrounds, compression speeds, and numbers of channels. Because floor and ceiling effects might have occurred for some listeners when SBR = –3 or 3, respectively, we decided to use only the scores obtained when SBR = 0 dB. This resulting measure is referred to as "overall intelligibility".

It was hypothesized that the intelligibility of speech in babble would worsen with increasing width of the auditory filters. This hypothesis was tested using a one-tailed test, since a specific direction of the effect was predicted. Figure 4 is a scatter plot of overall intelligibility versus overall filter widening. There was a small but significant negative correlation between the two measures ($r = -0.39$, $p = 0.043$, one-tailed), consistent with the hypothesis.

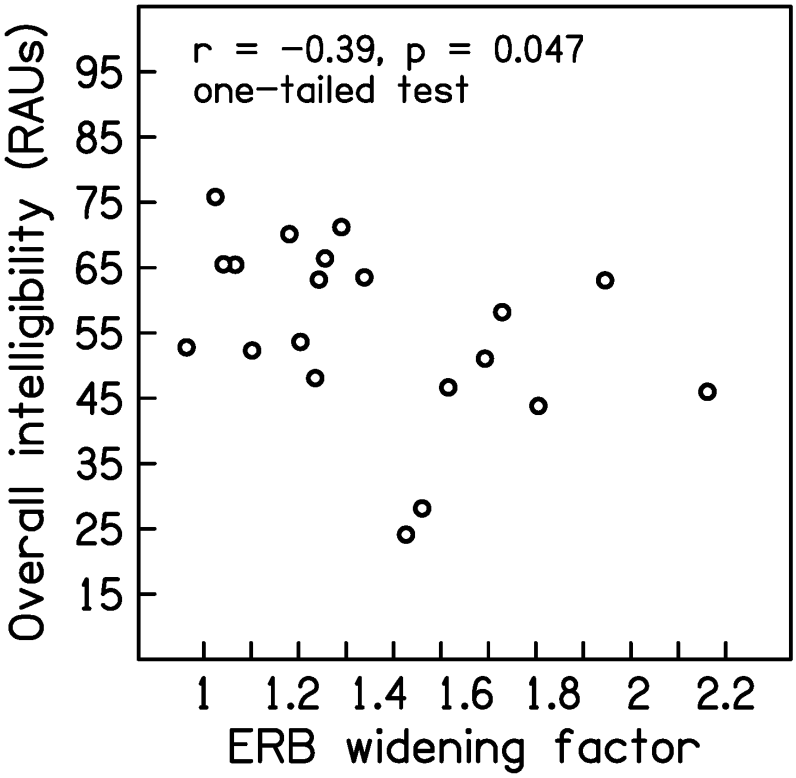

FIG. 4. Scatter plot showing the relationship between the overall widening of the auditory filters and overall intelligibility.

It was also hypothesized that overall intelligibility would improve with increasing sensitivity to TFS. Again, this hypothesis was assessed using a one-tailed test. The correlations between overall intelligibility and each of these measures of sensitivity to TFS were not significant (for TFS-AF $r = 0.40$, $p = 0.056$, one-tailed; and for the DLF $r = -0.09$, $p = 0.36$, one-tailed). Thus, the hypothesis was not supported.

2. Relationship between the effect of the number of channels and the psychoacoustic measures

Although there was no significant overall effect of number of channels, the number of channels did seem to have an effect for some listeners, but the effect varied across listeners.

To assess whether the individual effect of number of channels was related to any of the psychoacoustic measures, the intelligibility score in RAUs for each number of channels was averaged across backgrounds and compression speeds. The score obtained with 3 channels was subtracted from that obtained with 22 channels. Only the scores obtained for SBR = 0 were used in order to avoid floor or ceiling effects. The correlations between these difference scores and the results of the psychoacoustic tests were calculated. Two-tailed tests were performed as this analysis was exploratory.

The difference scores were not significantly correlated with the widening of the auditory filters (r = –0.20, p = 0.40, two-tailed), the DLFs (r = –0.15, p = 0.54, two-tailed), or the TFS-AF thresholds (r = –0.29, p = 0.25, two-tailed).

3. Relationship between the effect of compression speed and the psychoacoustic measures

It was hypothesized that better TFS processing abilities would be associated with a greater benefit from fast compression (Moore, 2008; Moore and Sek, 2016b). To assess this, the difference in performance between fast and slow compression was calculated by subtracting the average score in RAUs (across babble types) obtained with slow compression from that obtained with fast compression. Again, this difference was calculating using the scores for SBR = 0. Inspection of the scatter plots suggested that the relationships between difference scores and DLFs or TFS-AF thresholds were in the direction opposite to those expected. Thus, two-tailed tests were carried out. These indicated that the difference scores were not significantly correlated with the DLFs (r = 0.24, p = 0.34) or with the TFS-AF thresholds (r = –0.37, p = 0.129). Thus, the hypothesis was not supported.

IV. DISCUSSION

The results showed no effect of the number of compression channels or compression speed on the intelligibility of speech in 2-talker or 8-talker babble. This contrasts with some previous studies that reported significant effects of the number of compression channels

(Holube *et al.*, 2016; Alexander and Masterson, 2015; Yund and Buckles, 1995c; Bor *et al.*, 2008) and a main effect of the speed of compression or an interaction between the number of channels and the speed of compression (Alexander and Masterson, 2015; Holube *et al.*, 2016; Nábělek, 1983). There are several possible reasons for the discrepancies across studies.

Consider first the effect of number of channels. In most previous studies, the channels were used to shape the frequency response (Alexander and Masterson, 2015; Yund and Buckles, 1995a; Keidser and Grant, 2001; Crain and Yund, 1995), which was not the case in our study. When the number of channels is small, it is more difficult to achieve the target frequency response by adjusting the gain in the individual channels. Thus, in most previous studies the overall frequency response shape of the stimuli would have varied with the number of channels. This was noted by Yund and Buckles (1995c). They found that speech intelligibility improved with increasing number of channels from 4 to 8, but with no further increase at 16 channels. They analyzed the pattern of consonant confusions across conditions and found that most changes in performance could be accounted for by differences in the frequency response of the hearing aids; the gain at 4 kHz and above increased with increasing number of channels from 4 to 8 and from 8 to 16. They proposed that performance did not differ between 8 and 16 channels because the 8-channel hearing aid provided enough speech information to maximize performance. Our results suggest that when the target frequency response for medium-level speech is accurately achieved using a single filter applied before filtering into compression channels, there is no benefit of increasing the number of channels beyond 3.

Consider now the effect of compression speed. In many previous studies, the compression ratios were higher than used in our study, either because the hearing losses of the listeners were greater than here (Holube *et al.*, 2016) or because the compression ratios were set to arbitrary values, which often did not vary across frequency and were much higher than those prescribed for hearing aids (Náběłek, 1983; Plomp, 1994). Deleterious effects of compression are more likely to occur when the compression ratio is higher than needed to compensate for the loudness recruitment of the listener at any given frequency and when the compression is applied over a wide frequency range, especially when the compression is fast acting (Verschuure *et al.*, 1994; van Buuren *et al.*, 1999). Also, fast-acting compression applied at low frequencies can introduce waveform distortion, since the gain can change significantly within one cycle of sound (Moore *et al.*, 1999). Hence, the deleterious effects of fast-acting compression found in some previous studies might have been a consequence of the use of inappropriately high compression ratios, especially at low frequencies.

The median compression ratios used in the present work for channels with upper edges up to about 0.8-0.9 kHz were low, with medians of 1.01 for 3 channels, 1 and 1.67 for 6 channels, 1, 1.02, 1.05 and 1.24 for 12 channels, and 1, 1, 1.01, 1.05, 1.05, 1.11, 1.26, and 1.43 for 22 channels (Figures 1 and 2). When hearing loss occurs mainly at high frequencies, as is often the case (Agrawal *et al.*, 2008), the effect of clinically prescribed compression on the intelligibility of medium-level sounds might be small or absent. Similar points were made by De Gennaro *et al.* (1986) and Yund and Buckles (1995c), who hypothesized that the outcomes with multi-channel compression might depend on the degree of hearing loss. The work of Crain and Yund (1995) supports the idea that excessively high compression ratios might underlie the negative effects of multi-channel fast-acting compression that have sometimes been reported. Using fast-acting compression, they found that when the compression ratio was individually prescribed, there were no negative effects of increasing the number of channels from 1 to 31 on vowel or voiced stop-consonant discrimination. However, when the compression ratio was fixed across channels, there were negative effects of increasing the number of channels and of increasing the compression ratio, and an interaction between these two factors.

One study using individually prescribed compression ratios did show effects of compression speed. Alexander and Masterson (2015) compared fast and slow compression (release times of 40 and 640 ms, respectively), using 4, 8 or 16 channels (that were also used for frequency response shaping). They found that fast compression led to slightly better speech recognition than slow compression when the number of channels was 4, but the reverse was true when the number of channels was 16. However, best performance for both compression speeds was obtained with 8 channels. All of these effects were very small; the measure "generalized eta-squared" was 0.03 for the effect of number of channels, and 0.01 for the interaction of compression speed and number of channels. These small effects might have occurred because of changes in overall frequency response shape across conditions resulting from the use of the channels for frequency-response shaping as well as for compression.

Another possible reason for the discrepancy between our findings and those of some previous studies is related to the attack and release times used. Some of the release times used in previous work were shorter than those used here even for 'fast' compression (Holube *et al.*, 2016; Alexander and Masterson, 2015; Bor *et al.*, 2008; Yund and Buckles, 1995c). Short attack and release times might be associated with greater distortion of the temporal envelope and greater reductions in envelope modulation depth, depending on the design of the compressor (Stone and Moore, 2008). In studies where the attack and/or release times were varied systematically, very short time constants combined with high compression ratios usually led to worse performance (Stone and Moore, 2008; Nábělek, 1983) or decreased sound quality (Holube *et al.*, 2016). Our use of slightly longer time constants for the fast compressor, combined with the use of more appropriate compression ratios, would have avoided these deleterious effects.

Finally, the nature of the target speech and background used in the speech-recognition task might underlie differences across studies. Sentence material provides more contextual information than single words or nonsense syllables. Additionally, the background used here was natural speech (babble) and thus informational masking probably influenced performance, particularly for the two-talker babble (Brungart *et al.*, 2001; Hoen *et al.*, 2007). A small effect of compression speed or number of channels might be obscured by individual differences in cognitive and linguistic skills when listening to speech in speech (Brouwer *et al.*, 2012; Koelewijn *et al.*, 2012b; Newman *et al.*, 2015; Lecumberri and Cooke, 2006; Van Engen and Bradlow, 2007; Hoen *et al.*, 2007). Reducing contextual information in the target and linguistic content in the background might reduce this uncontrolled source of variability, but at the cost of reduced relevance to real-life scenarios. It is possible that effects of compression speed or the number of compression channels might be revealed even using sentence materials and babble backgrounds by assessing the outcome using measures of listening effort (Koelewijn *et al.*, 2012a; Koelewijn *et al.*, 2012b) rather than recognition scores. This remains to be determined.

One limitation of the present study is that the speech and babble were co-located and listening was monaural, which is not representative of real listening environments. Spatially separating speech and babble could reveal effects of the parameters of compression. For example, Moore *et al.* (2010a) assessed the effects of spatial separation and speed of compression on speech recognition in a 2-talker babble. They found that slow compression led to better performance than fast compression when the target and interference were spatially separated but not when they were co-located. However, the effect was small. The effect might have occurred because fast compression applied independently to the signal for each ear can alter interaural level differences (Wiggins and Seeber, 2011; Byrne and Noble, 1998; Hassager *et al.*, 2017). This effect can be avoided by synchronization of gain changes across ears (Kreisman *et al.*, 2010; Wiggins and Seeber, 2013; Hassager *et al.*, 2017).

The lack of interaction between speed of compression and type of background does not support the idea that fast compression with many channels improves “listening in the dips” of the masker, as suggested by Moore (2008). Our finding also contrasts with the finding of Moore *et al.* (1999) that fast compression with eight channels (but not four or less channels) improved the perception of speech in a modulated background sound, relative to linear amplification. However, there were differences across the studies in terms of task, stimuli, and background used. In particular, Moore *et al.* (1999) used a background that was highly modulated in amplitude and had large spectral dips, so as to maximize the role of dip listening. The small benefit found by them (less than a 1-dB improvement in the speech reception threshold) did not occur for more realistic background sounds. Our results are consistent with those of Alexander and Masterson (2015), who used steady and modulated background noises and found no interaction between the release time and the type of background.

It was of interest to assess whether variability in the effects of compression speed and number of channels across listeners was related to psychoacoustic measures. Overall performance in the sentence recognition task was negatively correlated with the widening of the auditory filters. This is consistent with the findings of Festen and Plomp (1983) and Moore *et al.* (1985).

Laurence *et al.* (1983) proposed that the amplitude-compression channels should be wider than the bandwidth of the auditory filters to avoid a reduction in spectral contrast with increasing number of channels. This proposal is supported by the findings of Croghan *et al.* (2014), who tested the preferences of hearing-impaired listeners for music signals subjected to linear amplification or multi-channel compression. They found that listeners with wider psychophysical tuning curves tended to prefer 3-channel over 18-channel compression, while listeners with sharper psychophysical tuning curves showed a small trend in the opposite direction. However, this applied only to classical music, and not to rock music. The results of Souza *et al.* (2012) are also consistent with the proposal of Laurence *et al.* (1983). Souza *et al.* (2012) showed that vowel identification was poorer for stimuli processed using 16-channel compression than for stimuli processed with linear amplification, and that this effect was greater for listeners with broader auditory filters. However, they used fast-acting compression with a fixed and rather high compression ratio of 3 in all channels. The deleterious effect of 16-channel compression would probably be smaller if listener-specific compression were used, and this would weaken the relationship between auditory-filter bandwidth and the effects of compression. Our data do not support the proposal of Laurence *et al.* (1983), since no correlation was found between overall filter bandwidths and changes in performance between 3 and 22 channels.

Finally, it was hypothesized that listeners with better TFS processing might benefit from fast compression, based on the suggestion of Moore (2008). This suggestion was weakly supported by the results of Moore and Sek (2016b), who found a small but significant correlation between the DLF at 2 kHz and the effect of compression speed on preference scores for speech and music: slow compression was preferred over fast compression more strongly with increasing DLF. However, no significant relationship was found between preference scores and TFS-AF thresholds. In the present study, no correlation was found between either of the TFS measures and the difference in performance between fast and slow compression. Thus, it appears that TFS processing is not related to changes in performance across compression speeds. This conclusion is consistent with the findings of Hopkins *et al.* (2012) for normal-hearing listeners listening to vocoded speech.

The psychoacoustic measures investigated here were not correlated with the effects of the number of channels or compression speed. The role of other factors, such as cognitive abilities, especially working memory, might be more significant (Gatehouse *et al.*, 2006a;

Gatehouse *et al.*, 2006b; Lunner and Sundewall-Thoren, 2007). However, although Reinhart and Souza (2016) found that slow compression led to better speech intelligibility than fast compression for reverberant speech using high compression ratios, no link between working memory and speed of compression was found.

The present work used realistic participant-specific compression ratios and showed no overall effect of compression speed. Hence, one cannot say that, in general, fast compression is better than slow compression, or vice versa. However, this does not rule out the possibility that individual listeners might achieve higher intelligibility with slow than with fast compression, or vice versa. It would be difficult in clinical practice to measure the intelligibility of speech in background sounds sufficiently accurately to demonstrate a reliable effect of compression speed for an individual. Hence, it seems reasonable to argue that the choice of compression speed for an individual should be determined based on subjective preferences and perhaps on measures of cognitive abilities such as working memory.

V. SUMMARY AND CONCLUSIONS

(1) The number of compression channels and the speed of compression had no effect on identification scores for sentences in 2-talker and 8-talker babble.

(2) There were no interactions between the parameters of the multi-channel compression and the type of background used.

(3) Overall speech scores were negatively correlated with the average widening of the auditory filters across 0.5, 1, and 2 kHz.

(4) None of the measures of TFS sensitivity were correlated with individual differences in performance between slow and fast compression.

(5) None of the psychoacoustic measures were correlated with individual differences in the effect of number of compression channels.

Acknowledgments

This work was supported by the H. B. Allen Trust and by the Engineering and Physical Sciences Research Council (UK) (grant RG78536). Author MAS was co-funded by the NIHR Manchester Biomedical Research Centre and Trust Charitable funds of the Central Manchester University Hospitals NHS Foundation Trust. We thank Borys Kowalewski and two anonymous reviewers for helpful comments on an earlier version of this paper.

Aazh, H., and Moore, B. C. J. (**2007**). "Dead regions in the cochlea at 4 kHz in elderly adults: relation to absolute threshold, steepness of audiogram, and pure tone average," J. Am. Acad. Audiol. **18**, 96-107.

Agrawal, Y., Platz, E. A., and Niparko, J. K. (**2008**). "Prevalence of hearing loss and differences by demographic characteristics among US adults: data from the National Health and Nutrition Examination Survey, 1999-2004," Arch. Intern. Med. **168**, 1522-1530.

Alexander, J. M., and Masterson, K. (**2015**). "Effects of WDRC release time and number of channels on output SNR and speech recognition," Ear Hear. **36**, e35-e49.

ANSI (**2014**). *ANSI S3.22-2014 - Specification of hearing aid characteristics* (American National Standards Institute, New York).

Baer, T., Moore, B. C. J., and Kluk, K. (**2002**). "Effects of lowpass filtering on the intelligibility of speech in noise for people with and without dead regions at high frequencies," J. Acoust. Soc. Am. **112**, 1133-1144.

Bor, S., Souza, P., and Wright, R. (**2008**). "Multichannel compression: effects of reduced spectral contrast on vowel identification," J. Speech. Lang. Hear. Res. **51**, 1315-1327.

Boyle, P. J., Nunn, T. B., O'Conner, A. F., and Moore, B. C. J. (**2013**). "STARR: A speech test for evaluation of the effectiveness of auditory prostheses under realistic conditions," Ear Hear. **34**, 203-212.

Bregman, A. S., Abramson, J., Doehring, P., and Darwin, C. J. (**1985**). "Spectral integration based on common amplitude modulation," Percept. Psychophys. **37**, 483-493.

British Society of Audiology (**2011**). *Pure tone air and bone conduction threshold audiometry with and without masking* (British Society of Audiology, Reading, UK).

Brouwer, S., Van Engen, K. J., Calandruccio, L., and Bradlow, A. R. (**2012**). "Linguistic contributions to speech-on-speech masking for native and non-native listeners: language familiarity and semantic content," J. Acoust. Soc. Am. **131**, 1449-1464.

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (**2001**). "Informational and energetic masking effects in the perception of multiple simultaneous talkers," J. Acoust. Soc. Am. **110**, 2527-2538.

Burkhard, M. D., and Sachs, R. M. (**1975**). "Anthropometric manikin for acoustic research," J. Acoust. Soc. Am. **58**, 214-222.

Bustamante, D. K., and Braida, L. D. (**1987**). "Multiband compression limiting for hearing-impaired listeners," J. Rehabil. Res. Dev. **24**, 149-160.

Byrne, D., Dillon, H., Ching, T., Katsch, R., and Keidser, G. (**2001**). "NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures," J. Am. Acad. Audiol. **12**, 37-51.

Byrne, D., and Noble, W. (**1998**). "Optimizing sound localization with hearing aids," Trends Amplif. **3**, 51-73.

Cornelisse, L. E., Seewald, R. C., and Jamieson, D. G. (**1995**). "The input/output formula: A theoretical approach to the fitting of personal amplification devices," J. Acoust. Soc. Am. **97**, 1854-1864.

Crain, T. R., and Yund, E. W. (**1995**). "The effect of multichannel compression on vowel and stop-consonant discrimination in normal-hearing and hearing-impaired subjects," Ear Hear. **16**, 529-543.

Croghan, N. B. H., Arehart, K. H., and Kates, J. M. (**2014**). "Music preferences with hearing aids: Effects of signal properties, compression settings, and listener characteristics," Ear Hear. **35**, e170-e184.

De Gennaro, S., Braida, L. D., and Durlach, N. I. (**1986**). "Multichannel syllabic compression for severely impaired listeners," J. Rehabil. Res. Dev. **23**, 17-24.

Ernst, S. M., and Moore, B. C. J. (**2013**). "Frequency difference limens at high frequencies for normal-hearing and hearing-impaired subjects," in *Proceedings of the International Conference on Acoustics - AIA-DAGA 2013*, Deutsche Gesellschaft für Akustik, Berlin, 1330-1331.

Festen, J. M., and Plomp, R. (**1983**). "Relations between auditory functions in impaired hearing," J. Acoust. Soc. Am. **73**, 652-662.

Füllgrabe, C., Harland, A. J., Sek, A. P., and Moore, B. C. J. (**2017**). "Development of a method for determining binaural sensitivity to temporal fine structure," Int. J. Audiol. **56**, 926-935.

Füllgrabe, C., and Moore, B. C. J. (**2017**). "Evaluation of a method for determining binaural sensitivity to temporal fine structure (TFS-AF Test) for older listeners with normal and impaired low-frequency hearing," Trends Hear. **21**, 1-14.

Gatehouse, S., Naylor, G., and Elberling, C. (**2006a**). "Linear and nonlinear hearing aid fittings - 1. Patterns of benefit," Int. J. Audiol. **45**, 130-152.

Gatehouse, S., Naylor, G., and Elberling, C. (**2006b**). "Linear and nonlinear hearing aid fittings - 2. Patterns of candidature," Int. J. Audiol. **45**, 153-171.

Glasberg, B. R., and Moore, B. C. J. (**1986**). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," J. Acoust. Soc. Am. **79**, 1020-1033.

Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.

Goldstein, J. L., and Srulovicz, P. (**1977**). "Auditory-nerve spike intervals as an adequate basis for aural frequency measurement," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans, and J. P. Wilson (Academic Press, London), pp. 337-346.

Hassager, H. G., Wiinberg, A., and Dau, T. (**2017**). "Effects of hearing-aid dynamic range compression on spatial perception in a reverberant environment," J. Acoust. Soc. Am. **141**, 2556-2568.

Hoen, M., Meunier, F., Grataloup, C.-L., Pellegrino, F., Grimault, N., Perrin, F., Perrot, X., Collet, L. (**2007**). "Phonetic and lexical interferences in informational masking during speech-in-speech comprehension," Speech Comm. **49**, 905-916.

Holube, I., Hamacher, V., and Killion, M. C. (**2016**). "Multi-channel compression: concepts and (early but timeless) results," Hear. Rev. **23**, 20-26.

Hopkins, K., King, A., and Moore, B. C. J. (**2012**). "The effect of compression speed on intelligibility: simulated hearing-aid processing with and without original temporal fine structure information," J. Acoust. Soc. Am. **132**, 1592-1601.

Hopkins, K., and Moore, B. C. J. (**2011**). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise," J. Acoust. Soc. Am. **130**, 334-349.

Kam, A. C., and Wong, L. L. (**1999**). "Comparison of performance with wide dynamic range compression and linear amplification," J. Am. Acad. Audiol. **10**, 445-57.

Kates, J. M. (**2010**). "Understanding compression: modeling the effects of dynamic-range compression in hearing aids," Int. J. Audiol. **49**, 395-409.

Keidser, G., Dillon, H., Flax, M., Ching, T., and Brewer, S. (**2011**). "The NAL-NL2 prescription procedure," Audiol. Res. **1:e24**, 88-90.

Keidser, G., and Grant, F. (**2001**). "The preferred number of channels (one, two, or four) in NAL-NL1 prescribed wide dynamic range compression (WDRC) devices," Ear Hear. **22**, 516-527.

Koelewijn, T., Zekveld, A. A., Festen, J. M., and Kramer, S. E. (**2012a**). "Pupil dilation uncovers extra listening effort in the presence of a single-talker masker," Ear Hear. **32**, 291-300.

Koelewijn, T., Zekveld, A. A., Festen, J. M., Rönnberg, J., and Kramer, S. E. (**2012b**). "Processing load induced by informational masking is related to linguistic abilities," Int. J. Otolaryngol. **2012**, 865731-865731.

Kreisman, B. M., Mazevski, A. G., Schum, D. J., and Sockalingam, R. (**2010**). "Improvements in speech understanding with wireless binaural broadband digital hearing instruments in adults with sensorineural hearing loss," Trends Amplif. **14**, 3-11.

Laurence, R. F., Moore, B. C. J., and Glasberg, B. R. (**1983**). "A comparison of behind-the-ear high-fidelity linear aids and two-channel compression hearing aids in the laboratory and in everyday life," Br. J. Audiol. **17**, 31-48.

Lecumberri, M. L. G., and Cooke, M. (**2006**). "Effect of masker type on native and non-native consonant perception in noise," J. Acoust. Soc. Am. **119**, 2445-2454.

Levitt, H., Cudahy, E., Hwang, W. H., Kennedy, E., and Link, C. (**1987**). "Towards a general measure of distortion," J. Rehabil. Res. Dev. **24**, 283-292.

Lippmann, R. P., Braida, L. D., and Durlach, N. I. (**1981**). "Study of multi-channel amplitude compression and linear amplification for persons with sensorineural hearing loss," J. Acoust. Soc. Am. **69**, 524-534.

Lunner, T., and Sundewall-Thoren, E. (**2007**). "Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid," J. Am. Acad. Audiol. **18**, 604-617.

Margolis, R. H., Eikelboom, R. H., Johnson, C., Ginter, S. M., Swanepoel, D. W., and Moore, B. C. J. (**2013**). "False air-bone gaps at 4 kHz in listeners with normal hearing and sensorineural hearing loss," Int. J. Audiol. **52**, 526-532.

Moore, B. C. J. (**1973**). "Frequency difference limens for narrow bands of noise," J. Acoust. Soc. Am. **54**, 888-896.

Moore, B. C. J. (**2001**). "Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids," Trends Amplif. **5**, 1-34.

Moore, B. C. J. (**2004**). "Dead regions in the cochlea: conceptual foundations, diagnosis, and clinical applications," Ear Hear. **25**, 98-116.

Moore, B. C. J. (**2007**). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues, 2nd Ed.* (Wiley, Chichester), pp. 1-332.

Moore, B. C. J. (**2008**). "The choice of compression speed in hearing aids: Theoretical and practical considerations, and the role of individual differences," Trends Amplif. **12**, 103-112.

Moore, B. C. J. (**2012**). "Effects of bandwidth, compression speed, and gain at high frequencies on preferences for amplified music," Trends Amplif. **16**, 159-172.

Moore, B. C. J., and Ernst, S. M. (**2012**). "Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code," J. Acoust. Soc. Am. **132**, 1542-1547.

Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (**2010a**). "Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task," J. Acoust. Soc. Am. **128**, 360-371.

Moore, B. C. J., Füllgrabe, C., and Stone, M. A. (**2011**). "Determination of preferred parameters for multi-channel compression using individually fitted simulated hearing aids and paired comparisons," Ear Hear. **32**, 556-568.

Moore, B. C. J., and Glasberg, B. R. (**1987**). "Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns," Hear. Res. **28**, 209-225.

Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (**2004**). "New version of the TEN test with calibrations in dB HL," Ear Hear. **25**, 478-487.

Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (**2010b**). "Development of a new method for deriving initial fittings for hearing aids with multi-channel compression: CAMEQ2-HF," Int. J. Audiol. **49**, 216-227.

Moore, B. C. J., Johnson, J. S., Clark, T. M., and Pluvinage, V. (**1992**). "Evaluation of a dual-channel full dynamic range compression system for people with sensorineural hearing loss," Ear Hear. **13**, 349-370.

Moore, B. C. J., Laurence, R. F., and Wright, D. (**1985**). "Improvements in speech intelligibility in quiet and in noise produced by two-channel compression hearing aids," Br. J. Audiol. **19**, 175-187.

Moore, B. C. J., Peters, R. W., and Stone, M. A. (**1999**). "Benefits of linear amplification and multi-channel compression for speech comprehension in backgrounds with spectral and temporal dips," J. Acoust. Soc. Am. **105**, 400-411.

Moore, B. C. J., and Popelka, G. R. (**2016**). "Introduction to hearing aids," in *Hearing Aids*, edited by G. R. Popelka, B. C. J. Moore, R. R. Fay, and A. N. Popper (Springer, New York), pp. 1-19.

Moore, B. C. J., and Sek, A. (**2016a**). "Comparison of the CAM2A and NAL-NL2 hearing-aid fitting methods for participants with a wide range of hearing losses," Int. J. Audiol. **55**, 93-100.

Moore, B. C. J., and Sek, A. (**2016b**). "Preferred compression speed for speech and music and its relationship to sensitivity to temporal fine structure," Trends Hear. **20**, 1-15.

Moore, B. C. J., and Shailer, M. J. (**1992**). "Modulation discrimination interference and auditory grouping," Philos. Trans. Royal Soc. B **336**, 339-346.

Moore, B. C. J., Stone, M. A., Füllgrabe, C., Glasberg, B. R., and Puria, S. (**2008**). "Spectro-temporal characteristics of speech at high frequencies, and the potential for restoration of audibility to people with mild-to-moderate hearing loss," Ear Hear. **29**, 907-922.

Moore, B. C. J., Vickers, D. A., and Mehta, A. (**2012**). "The effects of age on temporal fine structure sensitivity in monaural and binaural conditions," Int. J. Audiol. **51**, 715-721.

Mueller, G. R., Ricketts, T. A., and Bentler, R. (**2013**). *Modern Hearing Aids: Pre-fitting Testing and Considerations* (Plural, San Diego, CA), pp. 1-472.

Nábělek, I. V. (**1983**). "Performance of hearing-impaired listeners under various types of amplitude compression," J. Acoust. Soc. Am. **74**, 776-791.

Neuman, A. C., Bakke, M. H., Mackersie, C., and Hellman, S. (**1998**). "The effect of compression ratio and release time on the categorical rating of sound quality," J. Acoust. Soc. Am. **103**, 2273-2281.

Newman, R. S., Morini, G., Ahsan, F., and Kidd, G., Jr. (**2015**). "Linguistically-based informational masking in preschool children," J. Acoust. Soc. Am. **138**, EL93-EL98.

Patterson, R. D. (**1976**). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. **59**, 640-654.

Plomp, R. (**1988**). "The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function," J. Acoust. Soc. Am. **83**, 2322-2327.

Plomp, R. (**1994**). "Noise, amplification, and compression: Considerations of three main issues in hearing aid design," Ear Hear. **15**, 2-12.

Reinhart, P. N., and Souza, P. E. (**2016**). "Intelligibility and clarity of reverberant speech: Effects of wide dynamic range compression release time and working memory," J. Speech. Lang. Hear. Res. **59**, 1543-1554.

Rosen, S., Baker, R. J., and Kramer, S. (**1992**). "Characterizing changes in auditory filter bandwidth as a function of level," in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford), pp. 171-177.

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., Beaulac, S., Pumford, J. (**2005**). "The Desired Sensation Level multistage input/output algorithm," Trends Amplif. **9**, 159-197.

Sek, A., and Moore, B. C. J. (**2012**). "Implementation of two tests for measuring sensitivity to temporal fine structure," Int. J. Audiol. **51**, 58-63.

Sherbecoe, R. L., and Studebaker, G. A. (**2004**). "Supplementary formulas and table for calculating and interconverting speech recognition scores in transformed acrsine units," Int. J. Audiol. **43**, 442-448.

Souza, P., Wright, R., and Bor, S. (**2012**). "Consequences of broad auditory filters for identification of multichannel-compressed vowels," J. Speech. Lang. Hear. Res. **55**, 474-486.

Steinberg, J. C., and Gardner, M. B. (**1937**). "The dependence of hearing impairment on sound intensity," J. Acoust. Soc. Am. **9**, 11-23.

Stone, M. A., Füllgrabe, C., and Moore, B. C. J. (**2008**). "Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region," J. Acoust. Soc. Am. **124**, 2272-2282.

Stone, M. A., and Moore, B. C. J. (**1992**). "Syllabic compression: Effective compression ratios for signals modulated at different rates," Br. J. Audiol. **26**, 351-361.

Stone, M. A., and Moore, B. C. J. (**1999**). "Tolerable hearing-aid delays. I. Estimation of limits imposed by the auditory path alone using simulated hearing losses," Ear Hear. **20**, 182-192.

Stone, M. A., and Moore, B. C. J. (**2002**). "Tolerable hearing-aid delays. II. Estimation of limits imposed during speech production," Ear Hear. **23**, 325-338.

Stone, M. A., and Moore, B. C. J. (**2007**). "Quantifying the effects of fast-acting compression on the envelope of speech," J. Acoust. Soc. Am. **121**, 1654-1664.

Stone, M. A., and Moore, B. C. J. (**2008**). "Effects of spectro-temporal modulation changes produced by multi-channel compression on intelligibility in a competing-speech task," J. Acoust. Soc. Am. **123**, 1063-1076.

van Buuren, R. A., Festen, J., and Houtgast, T. (**1999**). "Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality," J. Acoust. Soc. Am. **105**, 2903-2913.

Van Engen, K. J., and Bradlow, A. R. (**2007**). "Sentence recognition in native- and foreign-language multi-talker background noise," J. Acoust. Soc. Am. **121**, 519-526.

Verschuure, J., Maas, A. J. J., Stikvoort, E., de Jong, R. M., Goedegebure, A., and Dreschler, W. A. (**1996**). "Compression and its effect on the speech signal," Ear Hear. **17**, 162-175.

Verschuure, J., Prinsen, T. T., and Dreschler, W. A. (**1994**). "The effect of syllabic compression and frequency shaping on speech intelligibility in hearing impaired people," Ear Hear. **15**, 13-21.

Vickers, D. A., Moore, B. C. J., and Baer, T. (**2001**). "Effects of lowpass filtering on the intelligibility of speech in quiet for people with and without dead regions at high frequencies," J. Acoust. Soc. Am. **110**, 1164-1175.

Villchur, E. (**1973**). "Signal processing to improve speech intelligibility in perceptive deafness," J. Acoust. Soc. Am. **53**, 1646-1657.

Vinay, and Moore, B. C. J. (**2007**). "Speech recognition as a function of highpass filter cutoff frequency for subjects with and without low-frequency cochlear dead regions," J. Acoust. Soc. Am. **122**, 542-553.

Wiggins, I. M., and Seeber, B. U. (**2011**). "Dynamic-range compression affects the lateral position of sounds," J. Acoust. Soc. Am. **130**, 3939-3953.

Wiggins, I. M., and Seeber, B. U. (**2013**). "Linking dynamic-range compression across the ears can improve speech intelligibility in spatially separated noise," J. Acoust. Soc. Am. **133**, 1004-1016.

Woods, W. S., Van Tasell, D. J., Rickert, M. E., and Trine, T. D. (**2006**). "SII and fit-to-target analysis of compression system performance as a function of number of compression channels," Int. J. Audiol. **45**, 630-644.

Yund, E. W., and Buckles, K. M. (**1995a**). "Discrimination of multi-channel compressed speech in noise: Long-term learning in hearing-impaired subjects," Ear Hear. **16**, 417-427.

Yund, E. W., and Buckles, K. M. (**1995b**). "Enhanced speech perception at low signal-to-noise ratios with multichannel compression hearing aids," J. Acoust. Soc. Am. **97**, 1224-1240.

Yund, E. W., and Buckles, K. M. (**1995c**). "Multichannel compression hearing aids: Effect of number of channels on speech discrimination in noise," J. Acoust. Soc. Am. **97**, 1206-1223.

Yund, E. W., Simon, H. J., and Efron, R. (**1987**). "Speech discrimination with an 8-channel compression hearing aid and conventional aids in background of speech-band noise," J. Rehabil. Res. Dev. **24**, 160-80.